MLLNVLRICI	IVCLVNDGAG	KHSEGRERTK	TYSLNSRGYF	40
RKERGARRSK	ILLVNTKGLD	EPHIGHGDFG	LVAELFDSTR	80
THTNRKEPDM	NKVKLFSTVA	HGNKSARRKA	YNGSRRNIFS	120
RRSFDKRNTE	VTEKPGAKMF	WNNFLVKMNG	APQNTSHGSK	160
AQEIMKEACK	TLPFTQNIVH	ENCDRMVIQN	NLCFGKCISL	200
HVPNQQDRRN	TCSHCLPSKF	TLNHLTLNCT	GSKNVVKVVM	240
MVEECTCEAH	KSNFHQTAQF	NMDTSTTLHH		270

Figure 1. Deduced amino acid sequence of Xenopus cerberus protein. SEQ ID NO:1.

Figure 2. Nucleotide sequence of the full-length cerberus DNA derived from the Xenopus organizer. The sense strand is on top (in the 5' to 3' direction) and the antisense strand on the bottom line (on the opposite direction). SEQ ID NO:2.

```
GAATTCCCAG CAAGTCGCTC AGAAACACTG CAGGGTCTAG ATATCATACA ATGTTACTAA     60
CTTAAGGGTC GTTCAGCGAG TCTTTGTGAC GTCCCAGATC TATAGTATGT TACAATGATT

ATGTACTCAG GATCTGTATT ATCGTCTGCC TTGTGAATGA TGGAGCAGGA AAACACTCAG    120
TACATGAGTC CTAGACATAA TAGCAGACGG AACACTTACT ACCTCGTCCT TTTGTGAGTC

AAGGACGAGA AAGGACAAAA ACATATTCAC TTAACAGCAG AGGTTACTTC AGAAAAGAAA    180
TTCCTGCTCT TTCCTGTTTT TGTATAAGTG AATTGTCGTC TCCAATGAAG TCTTTTCTTT

GAGGAGCACG TAGGAGCAAG ATTCTGCTGG TGAATACTAA AGGTCTTGAT GAACCCCACA    240
CTCCTCGTGC ATCCTCGTTC TAAGACGACC ACTTATGATT TCCAGAACTA CTTGGGGTGT

TTGGGCATGG TGATTTTCGC TTAGTAGCTG AACTATTTGA TTCCACCAGA ACACATACAA    300
AACCCGTACC ACTAAAAGCG AATCATCGAC TTGATAAACT AAGGTGGTCT TGTGTATGTT

ACAGAAAAGA GCCAGACATG AACAAAGTCA AGCTTTTCTC AACAGTTGCC CATGGAAACA    360
TGTCTTTTCT CGGTCTGTAC TTGTTTCAGT TCGAAAAGAG TTGTCAACGG GTACCTTTGT

AAAGTGCAAG AAGAAAAGCT TACAATGGTT CTAGAAGGAA TATTTTTCCT CGCCGTTCTT    420
TTTCACGTTC TTCTTTTCGA ATGTTACCAA GATCTTCCTT ATAAAAAGGA GCGGCAAGAA

TTGATAAAAG AAATACAGAG GTTACTGAAA AGCCTGGTGC CAAGATGTTC TGGAACAATT    480
AACTATTTTC TTTATGTCTC CAATGACTTT TCGGACCACG GTTCTACAAG ACCTTGTTAA

TTTTGGTTAA AATGAATGGA GCCCCACAGA ATACAAGCCA TGGCAGTAAA GCACAGGAAA    540
AAAACCAATT TTACTTACCT CGGGGTGTCT TATGTTCGGT ACCGTCATTT CGTGTCCTTT

TAATGAAAGA AGCTTGCAAA ACCTTGTTTT TCACTCAGAA TATTGTACAT GAAAACTGTG    600
ATTACTTTCT TCGAACGTTT TGGAACAAAA AGTGAGTCTT ATAACATGTA CTTTTGACAC

ACAGGATGGT GATACAGAAC AATCTGTGCT TTGGTAAATG CATCTCTCTC CATGTTCCAA    660
TGTCCTACCA CTATGTCTTG TTAGACACGA AACCATTTAC GTAGAGAGAG GTACAAGGTT

ATCAGCAAGA TCGACGAAAT ACTTGTTCCC ATTGCTTGCC GTCCAAATTT ACCCTGAACC    720
TAGTCGTTCT AGCTGCTTTA TGAACAAGGG TAACGAACGG CAGGTTTAAA TGGGACTTGG

ACCTGACGCT GAATTGTACT GGATCTAAGA ATGTAGTAAA GGTTGTCATG ATGGTAGAGG    780
TGGACTGCGA CTTAACATGA CCTAGATTCT TACATCATTT CCAACAGTAC TACCATCTCC

AATGCACGTG TGAAGCTCAT AAGAGCAACT TCCACCAAAC TGCACAGTTT AACATGGATA    840
TTACGTGCAC ACTTCGAGTA TTCTCGTTGA AGGTGGTTTG ACGTGTCAAA TTGTACCTAT

CATCTACTAC CCTGCACCAT TAAAGGACTG CCATACAGTA TGGAAATGCC CTTTTGTTGG    900
GTAGATGATG GGACGTGGTA ATTTCCTGAC GGTATGTCAT ACCTTTACGG GAAAACAACC

AATATTTGTT ACATACTATG CATCTAAAGC ATTATGTTGC CTTCTATTTC ATATAACCAC    960
TTATAAACAA TGTATGATAC GTAGATTTCG TAATACAACG GAAGATAAAG TATATTGGTG

ATGGAATAAG GATTGTATGA ATTATAATTA ACAAATGGCA TTTTGTGTAA CATGCAAGAT   1020
TACCTTATTC CTAACATACT TAATATTAAT TGTTTACCGT AAAACACATT GTACGTTCTA
```

CTCTGTTCCA TCAGTTGCAA GATAAAAGGC AATATTTGTT TGACTTTTTT TCTACAAAAT 1080
GAGACAAGGT AGTCAACGTT CTATTTTCCG TTATAAACAA ACTGAAAAAA AGATGTTTTA

GAATACCCAA ATATATGATA AGATAATGGG GTCAAAACTG TTAAGGGGTA ATGTAATAAT 1140
CTTATGGGTT TATATACTAT TCTATTACCC CAGTTTTGAC AATTCCCCAT TACATTATTA

AGGGACTAAG TTTGCCCAGG AGCAGTGACC CATAACAACC AATCAGCAGG TATGATTTAC 1200
TCCCTGATTC AAACGGGTCC TCGTCACTGG GTATTGTTGG TTAGTCGTCC ATACTAAATG

TGGTCACCTG TTTAAAAGCA AACATCTTAT TGGTTGCTAT GGGTTACTGC TTCTGGGCAA 1260
ACCAGTGGAC AAATTTTCGT TTGTAGAATA ACCAACGATA CCCAATGACG AAGACCCGTT

AATGTGTGCC TCATAGGGGG GTTAGTGTGT TGTGTACTGA ATAAATTGTA TTTATTTCAT 1320
TTACACACGG AGTATCCCCC CAATCACACA ACACATGACT TATTTAACAT AAATAAAGTA

TGTTACAAAA AAAAAAAA
ACAATGTTTT TTTTTTTT

Fig. 2. (Continuation page 2, SEQ ID NO:2).

MSRTRKVDSL LLLAIPGLAL LLLPNAYCAS CEPVRIPMCK SMPWNMTKMP NHLHHSTQAN 60

AILAIEQFEG LLTTECSQDL LFFLCAMYAP ICTIDFQHEP IKPCKSVCER ARAGCEPILI 120

KYRHTWPESL ACEELPVYDR GVCISPEAIV TVEQGTDSMP DFSMDSNNGN CGSGREHCKC 180

KPMKATQKTY LKNNYNYVIR AKVKEVKVKC HDATAIVEVK EILKSSLVNI PKDTVTLYTN 240

SGCLCPQLVA NEEYIIMGYE DKERTRLLLV EGSLAEKWRD RLAKKVKRWD QKLRRPRKSK 300

DPVAPIPNKN SNSRQARS

Figure 3. Deduced amino acid sequence of Xenopus frazzled protein. SEQ ID NO:3.

Figure 4. Nucleotide sequence of the full-length frazzled cDNA derived from the Xenopus organizer. The sense strand of the DNA on top (5' to 3' direction) and the antisense strand on the bottom line (opposite direction). SEQ ID NO:4.

```
GAATTCCCTT TCACACAGGA CTCCTGGCAG AGGTGAATGG TTAGCCCTAT GGATTTGGTT     60
CTTAAGGGAA AGTGTGTCCT GAGGACCGTC TCCACTTACC AATCGGGATA CCTAAACCAA

TGTTGATTTT GACACATGAT TGATTGCTTT CAGATAGGAT TGAAGGACTT GGATTTTTAT    120
ACAACTAAAA CTGTGTACTA ACTAACGAAA GTCTATCCTA ACTTCCTGAA CCTAAAAATA

CTAATTCTGC ACTTTTAAAT TATCTGAGTA ATTGTTCATT TTGTATTGGA TGGGACTAAA    180
GATTAAGACG TGAAAATTTA ATAGACTCAT TAACAAGTAA AACATAACCT ACCCTGATTT

GATAAACTTA ACTCCTTGCT TTTGACTTGC CCATAAACTA TAAGGTGGGG TGAGTTGTAG    240
CTATTTGAAT TGAGGAACGA AAACTGAACG GGTATTTGAT ATTCCACCCC ACTCAACATC

TTGCTTTTAC ATGTGCCCAG ATTTTCCCTG TATTCCCTGT ATTCCCTCTA AAGTAAGCCT    300
AACGAAAATG TACACGGGTC TAAAAGGGAC ATAAGGGACA TAAGGGAGAT TTCATTCGGA

ACACATACAG GTTGGGCAGA ATAACAATGT CTCGAACAAG GAAAGTGGAC TCATTACTGC    360
TGTGTATGTC CAACCCGTCT TATTGTTACA GAGCTTGTTC CTTTCACCTG AGTAATGACG

TACTGGCCAT ACCTGGACTG GCGCTTCTCT TATTACCCAA TGCTTACTGT GCTTCGTGTG    420
ATGACCGGTA TGGACCTGAC CGCGAAGAGA ATAATGGGTT ACGAATGACA CGAAGCACAC

AGCCTGTGCG GATCCCCATG TGCAAATCTA TGCCATGGAA CATGACCAAG ATGCCCAACC    480
TCGGACACGC CTAGGGGTAC ACGTTTAGAT ACGGTACCTT GTACTGGTTC TACGGGTTGG

ATCTCCACCA CAGCACTCAA GCCAATGCCA TCCTGGCAAT TGAACAGTTT GAAGGTTTGC    540
TAGAGGTGGT GTCGTGAGTT CGGTTACGGT AGGACCGTTA ACTTGTCAAA CTTCCAAACG

TGACCACTGA ATGTAGCCAG GACCTTTTGT TCTTTCTGTG TGCCATGTAT GCCCCCATTT    600
ACTGGTGACT TACATCGGTC CTGGAAAACA AGAAAGACAC ACGGTACATA CGGGGGTAAA

GTACCATCGA TTTCCAGCAT GAACCAATTA AGCCTTGCAA GTCCGTGTGC GAAAGGGCCA    660
CATGGTAGCT AAAGGTCGTA CTTGGTTAAT TCGGAACGTT CAGGCACACG CTTTCCCGGT

GGGCCGGCTG TGAGCCCATT CTCATAAAGT ACCGGCACAC TTGGCCAGAG AGCCTGGCAT    720
CCCGGCCGAC ACTCGGGTAA GAGTATTTCA TGGCCGTGTG AACCGGTCTC TCGGACCGTA

GTGAAGAGCT GCCCGTATAT GACAGAGGAG TCTGCATCTC CCCAGAGGCT ATCGTCACAG    780
CACTTCTCGA CGGGCATATA CTGTCTCCTC AGACGTAGAG GGGTCTCCGA TAGCAGTGTC

TGGAACAAGG AACAGATTCA ATGCCAGACT TCTCCATGGA TTCAAACAAT GGAAATTGCG    840
ACCTTGTTCC TTGTCTAAGT TACGGTCTGA AGAGGTACCT AAGTTTGTTA CCTTTAACGC

GAAGCGGCAG GGAGCACTGT AAATGCAAGC CCATGAAGGC AACCCAAAAG ACGTATCTCA    900
CTTCGCCGTC CCTCGTGACA TTTACGTTCG GGTACTTCCG TTGGGTTTTC TGCATAGAGT

AGAATAATTA CAATTATGTA ATCAGAGCAA AAGTGAAAGA GGTGAAAGTG AAATGCCACG    960
TCTTATTAAT GTTAATACAT TAGTCTCGTT TTCACTTTCT CCACTTTCAC TTTACGGTGC

ACGCAACAGC AATTGTGGAA GTAAAGGAGA TTCTCAAGTC TTCCCTAGTG AACATTCCTA   1020
TGCGTTGTCG TTAACACCTT CATTTCCTCT AAGAGTTCAG AAGGGATCAC TTGTAAGGAT
```

AAGACACAGT GACACTGTAC ACCAACTCAG GCTGCTTGTG CCCCCAGCTT GTTGCCAATG 1080
TTCTGTGTCA CTGTGACATG TGGTTGAGTC CGACGAACAC GGGGGTCGAA CAACGGTTAC

AGGAATACAT AATTATGGGC TATGAAGACA AAGAGCGTAC CAGGCTTCTA CTAGTGGAAG 1140
TCCTTATGTA TTAATACCCG ATACTTCTGT TTCTCGCATG GTCCGAAGAT GATCACCTTC

GATCCTTGGC CGAAAAATGG AGAGATCGTC TTGCTAAGAA AGTCAAGCGC TGGGATCAAA 1200
CTAGGAACCG GCTTTTTACC TCTCTAGCAG AACGATTCTT TCAGTTCGCG ACCCTAGTTT

AGCTTCGACG TCCCAGGAAA AGCAAAGACC CCGTGGCTCC AATTCCCAAC AAAAACAGCA 1260
TCGAAGCTGC AGGGTCCTTT TCGTTTCTGG GGCACCGAGG TTAAGGGTTG TTTTTGTCGT

ATTCCAGACA AGCGCGTAGT TAGACTAACG GAAAGGTGTA TGGAAACTCT ATGGACTTTG 1320
TAAGGTCTGT TCGCGCATCA ATCTGATTGC CTTTCCACAT ACCTTTGAGA TACCTGAAAC

AAACTAAGAT TTGCATTGTT GGAAGAGCAA AAAAGAAATT GCACTACAGC ACGTTATATT 1380
TTTGATTCTA AACGTAACAA CCTTCTCGTT TTTTCTTTAA CGTGATGTCG TGCAATATAA

CTATTGTTTA CTACAAGAAG CTGGTTTAGT TGATTGTAGT TCTCCTTTCC TTCTTTTTTT 1440
GATAACAAAT GATGTTCTTC GACCAAATCA ACTAACATCA AGAGGAAAGG AAGAAAAAAA

TTATAACTAT ATTTGCACGT GTTCCCAGGC AATTGTTTTA TTCAACTTCC AGTGACAGAG 1500
AATATTGATA TAAACGTGCA CAAGGGTCCG TTAACAAAAT AAGTTGAAGG TCACTGTCTC

CAGTGACTGA ATGTCTCAGC CTAAAGAAGC TCAATTCATT TCTGATCAAC TAATGGTGAC 1560
GTCACTGACT TACAGAGTCG GATTTCTTCG AGTTAAGTAA AGACTAGTTG ATTACCACTG

AAGTGTTTGA TACTTGGGGA AAGTGAACTA ATTGCAATGG TAAATCAGAG AAAAGTTGAC 1620
TTCACAAACT ATGAACCCCT TTCACTTGAT TAACGTTACC ATTTAGTCTC TTTTCAACTG

CAATGTTGCT TTTCCTGTAG ATGAACAAGT GAGAGATCAC ATTTAAATGA TGATCACTTT 1680
GTTACAACGA AAAGGACATC TACTTGTTCA CTCTCTAGTG TAAATTTACT ACTAGTGAAA

CCATTTAATA CTTTCAGCAG TTTTAGTTAG ATGACATGTA GGATGCACCT AAATCTAAAT 1740
GGTAAATTAT GAAAGTCGTC AAAATCAATC TACTGTACAT CCTACGTGGA TTTAGATTTA

ATTTTATCAT AAATGAAGAG CTGGTTTAGA CTGTATGGTC ACTGTTGGGA AGGTAAATGC 1800
TAAAATAGTA TTTACTTCTC GACCAAATCT GACATACCAG TGACAACCCT TCCATTTACG

CTACTTTGTC AATTCTGTTT TAAAAATTGC CTAAATAAAT ATTAAGTCCT AAATAAAAAA 1860
GATGAAACAG TTAAGACAAA ATTTTTAACG GATTTATTTA TAATTCAGGA TTTATTTTTT

AAAAAAAAAA AAAAA
TTTTTTTTTT TTTTT

Fig. 4. (Continuation page 2, SEQ ID NO:4).

MLLLFRAIPM LLLGLMVLQT DCEIAQYYID EEEPPGTVIA VLSQHSIFNT TDIPATNFRL 60

MKQFNNSLIG VRESDGQLSI MERIDREQIC RQSLHCNLAL DVVSFSKGHF KLLNVKVEVR 120

DINDHSPHFP SEIMHVEVSE SSSVGTRIPL EIAIDEDVGS NSIQNFQISN NSHFSIDVLT 180

RADGVKYADL VLMRELDREI QPTYIMELLA MDGGVPSLSG TAVVNIRVLD FNDNSPVFER 240

STIAVDLVED APLGYLLLEL HATDDDEGVN GEIVYGFSTL ASQEVRQLFK INSRTGSVTL 300

EGQVDFETKQ TYEFEVQAQD LGPNPLTATC KVTVHILDVN DNTPAITITP LTTVNAGVAY 360

IPETATKENF IALISTTDRA SGSNGQVRCT LYGHEHFKLQ QAYEDSYMIV TTSTLDRENI 420

AAYSLTVVAE DLGFPSLKTK KYYTVKVSDE NDNAPVFSKP QYEASILENN APGSYITTVI 480

ARDSDSDQNG KVNYRLVDAK VMGQSLTTFV SLDADSGVLR AVRSLDYEKL KQLDFEIEAA 540

DNGIPQLSTR VQLNLRIVDQ NDNCPVITNP LLNNGSGEVL LPISAPQNYL VFQLKAEDSD 600

EGHNSQLFYT ILRDPSRLFA INKESGEVFL KKQLNSDHSE DLSIVVAVYD LGRPSLSTNA 660

TVKFILTDSF PSNVEVVILQ PSAEEQHQID MSIIFIAVLA GGCALLLLAI FFVACTCKKK 720

AGEFKQVPEQ HGTCNEERLL STPSPQSVSS SLSQSESCQL SINTESENCS VSSNQEQHQQ 780

TGIKHSISVP SYHTSGWHLD NCAMSISGHS HMGHISTKVQ WAKEIVTSMT VTLILVENQK 840

RRALSSQCRH KPVLNTQMNQ QGSDMPITIS ATESTRVQKM GTAHCNMKRA IDCLTL

Figure 5. Deduced amino acid sequence of the Xenopus PAPC (paraxial protocadherin) protein. It encodes a member of the cadherin family of transmembrane proteins that has dorsalizing activity when constructs are injected into Xenopus embryos. SEQ ID NO:5.

Figure 6. Nucleotide sequence of the full-length PAPC cDNA derived from the Xenopus organizer. The sense strand of the DNA is shown in the top line (in the 5' to 3' direction), and the bottom line shows the antisense strand (opposite orientation). SEQ ID NO:6.

```
GAATTCCCAG AGATGAACTC CTTGAGATTG TTTTAAATGA CTGCAGGTCT GGAAGGATTC     60
CTTAAGGGTC TCTACTTGAG GAACTCTAAC AAAATTTACT GACGTCCAGA CCTTCCTAAG

ACATTGCCAC ACTGTTTCTA GGCATGAAAA AACTGCAAGT TTCAACTTTG TTTTTGGTGC    120
TGTAACGGTG TGACAAAGAT CCGTACTTTT TTGACGTTCA AAGTTGAAAC AAAAACCACG

AACTTTGATT CTTCAAGATG CTGCTTCTCT TCAGAGCCAT TCCAATGCTG CTGTTGGGAC    180
TTGAAACTAA GAAGTTCTAC GACGAAGAGA AGTCTCGGTA AGGTTACGAC GACAACCCTG

TGATGGTTTT ACAAACAGAC TGTGAAATTG CCCAGTACTA CATAGATGAA GAAGAACCCC    240
ACTACCAAAA TGTTTGTCTG ACACTTTAAC GGGTCATGAT GTATCTACTT CTTCTTGGGG

CTGGCACTGT AATTGCAGTG TTGTCACAAC ACTCCATATT TAACACTACA GATATACCTG    300
GACCGTGACA TTAACGTCAC AACAGTGTTG TGAGGTATAA ATTGTGATGT CTATATGGAC

CAACCAATTT CCGTCTAATG AAGCAATTTA ATAATTCCCT TATCGGAGTC CGTGAGAGTG    360
GTTGGTTAAA GGCAGATTAC TTCGTTAAAT TATTAAGGGA ATAGCCTCAG GCACTCTCAC

ATGGGCAGCT GAGCATCATG GAGAGGATTG ACCGGGAGCA AATCTGCAGG CAGTCCCTTC    420
TACCCGTCGA CTCGTAGTAC CTCTCCTAAC TGGCCCTCGT TTAGACGTCC GTCAGGGAAG

ACTGCAACCT GGCTTTGGAT GTGGTCAGCT TTTCCAAAGG ACACTTCAAG CTTCTGAACG    480
TGACGTTGGA CCGAAACCTA CACCAGTCGA AAAGGTTTCC TGTGAAGTTC GAAGACTTGC

TGAAAGTGGA GGTGAGAGAC ATTAATGACC ATAGCCCTCA CTTTCCCAGT GAAATAATGC    540
ACTTTCACCT CCACTCTCTG TAATTACTGG TATCGGGAGT GAAAGGGTCA CTTTATTACG

ATGTGGAGGT GTCTGAAAGT TCCTCTGTGG GCACCAGGAT TCCTTTAGAA ATTGCAATAG    600
TACACCTCCA CAGACTTTCA AGGAGACACC CGTGGTCCTA AGGAAATCTT TAACGTTATC

ATGAAGATGT TGGGTCCAAC TCCATCCAGA ACTTTCAGAT CTCAAATAAT AGCCACTTCA    660
TACTTCTACA ACCCAGGTTG AGGTAGGTCT TGAAAGTCTA GAGTTTATTA TCGGTGAAGT

GCATTGATGT GCTAACCAGA GCAGATGGGG TGAAATATGC AGATTTAGTC TTAATGAGAG    720
CGTAACTACA CGATTGGTCT CGTCTACCCC ACTTTATACG TCTAAATCAG AATTACTCTC

AACTGGACAG GGAAATCCAG CCAACATACA TAATGGAGCT ACTAGCAATG GATGGGGGTG    780
TTGACCTGTC CCTTTAGGTC GGTTGTATGT ATTACCTCGA TGATCGTTAC CTACCCCCAC

TACCATCACT ATCTGGTACT GCAGTGGTTA ACATCCGAGT CCTGGACTTT AATGATAACA    840
ATGGTAGTGA TAGACCATGA CGTCACCAAT TGTAGGCTCA GGACCTGAAA TTACTATTGT

GCCCAGTGTT TGAGAGAAGC ACCATTGCTG TGGACCTAGT AGAGGATGCT CCTCTGGGAT    900
CGGGTCACAA ACTCTCTTCG TGGTAACGAC ACCTGGATCA TCTCCTACGA GGAGACCCTA

ACCTTTTGTT GGAGTTACAT GCTACTGACG ATGATGAAGG AGTGAATGGA GAAATTGTTT    960
TGGAAAACAA CCTCAATGTA CGATGACTGC TACTACTTCC TCACTTACCT CTTTAACAAA

ATGGATTCAG CACTTTGGCA TCTCAAGAGG TACGTCAGCT ATTTAAAATT AACTCCAGAA   1020
TACCTAAGTC GTGAAACCGT AGAGTTCTCC ATGCAGTCGA TAAATTTTAA TTGAGGTCTT
```

```
CTGGCAGTGT TACTCTTGAA GGCCAAGTTG ATTTTGAGAC CAAGCAGACT TACGAATTTG   1080
GACCGTCACA ATGAGAACTT CCGGTTCAAC TAAAACTCTG GTTCGTCTGA ATGCTTAAAC

AGGTACAAGC CCAAGATTTG GGCCCCAACC CACTGACTGC TACTTGTAAA GTAACTGTTC   1140
TCCATGTTCG GGTTCTAAAC CCGGGGTTGG GTGACTGACG ATGAACATTT CATTGACAAG

ATATACTTGA TGTAAATGAT AATACCCCAG CCATCACTAT TACCCCTCTG ACTACTGTAA   1200
TATATGAACT ACATTTACTA TTATGGGGTC GGTAGTGATA ATGGGGAGAC TGATGACATT

ATGCAGGAGT TGCCTATATT CCAGAAACAG CCACAAAGGA GAACTTTATA GCTCTGATCA   1260
TACGTCCTCA ACGGATATAA GGTCTTTGTC GGTGTTTCCT CTTGAAATAT CGAGACTAGT

GCACTACTGA CAGAGCCTCT GGATCTAATG GACAAGTTCG CTGTACTCTT TATGGACATG   1320
CGTGATGACT GTCTCGGAGA CCTAGATTAC CTGTTCAAGC GACATGAGAA ATACCTGTAC

AGCACTTTAA ACTACAGCAA GCTTATGAGG ACAGTTACAT GATAGTTACC ACCTCTACTT   1380
TCGTGAAATT TGATGTCGTT CGAATACTCC TGTCAATGTA CTATCAATGG TGGAGATGAA

TAGACAGGGA AAACATAGCA GCGTACTCTT TGACAGTAGT TGCAGAAGAC CTTGGCTTCC   1440
ATCTGTCCCT TTTGTATCGT CGCATGAGAA ACTGTCATCA ACGTCTTCTG GAACCGAAGG

CCTCATTGAA GACCAAAAAG TACTACACAG TCAAGGTTAG TGATGAGAAT GACAATGCAC   1500
GGAGTAACTT CTGGTTTTTC ATGATGTGTC AGTTCCAATC ACTACTCTTA CTGTTACGTG

CTGTATTTTC TAAACCCCAG TATGAAGCTT CTATTCTGGA AAATAATGCT CCAGGCTCTT   1560
GACATAAAAG ATTTGGGGTC ATACTTCGAA GATAAGACCT TTTATTACGA GGTCCGAGAA

ATATAACTAC AGTGATAGCC AGAGACTCTG ATAGTGATCA AAATGGCAAA GTAAATTACA   1620
TATATTGATG TCACTATCGG TCTCTGAGAC TATCACTAGT TTTACCGTTT CATTTAATGT

GACTTGTGGA TGCAAAAGTG ATGGGCCAGT CACTAACAAC ATTTGTTTCT CTTGATGCGG   1680
CTGAACACCT ACGTTTTCAC TACCCGGTCA GTGATTGTTG TAAACAAAGA GAACTACGCC

ACTCTGGAGT ATTGAGAGCT GTTAGGTCTT TAGACTATGA AAAACTTAAA CAACTGGATT   1740
TGAGACCTCA TAACTCTCGA CAATCCAGAA ATCTGATACT TTTTGAATTT GTTGACCTAA

TTGAAATTGA AGCTGCAGAC AATGGGATCC CTCAACTCTC CACTCGCGTT CAACTAAATC   1800
AACTTTAACT TCGACGTCTG TTACCCTAGG GAGTTGAGAG GTGAGCGCAA GTTGATTTAG

TCAGAATAGT TGATCAAAAT GATAATTGCC CTGTGATAAC TAATCCTCTT CTTAATAATG   1860
AGTCTTATCA ACTAGTTTTA CTATTAACGG GACACTATTG ATTAGGAGAA GAATTATTAC

GCTCGGGTGA AGTTCTGCTT CCCATCAGCG CTCCTCAAAA CTATTTAGTT TTCCAGCTCA   1920
CGAGCCCACT TCAAGACGAA GGGTAGTCGC GAGGAGTTTT GATAAATCAA AAGGTCGAGT

AAGCCGAGGA TTCAGATGAA GGGCACAACT CCCAGCTGTT CTATACCATA CTGAGAGATC   1980
TTCGGCTCCT AAGTCTACTT CCCGTGTTGA GGGTCGACAA GATATGGTAT GACTCTCTAG

CAAGCAGATT GTTTGCCATT AACAAAGAAA GTGGTGAAGT GTTCCTGAAA AAACAATTAA   2040
GTTCGTCTAA CAAACGGTAA TTGTTTCTTT CACCACTTCA CAAGGACTTT TTTGTTAATT

ACTCTGACCA TTCAGAGGAC TTGAGCATAG TAGTTGCAGT GTATGACTTG GGAAGACCTT   2100
TGAGACTGGT AAGTCTCCTG AACTCGTATC ATCAACGTCA CATACTGAAC CCTTCTGGAA

CATTATCCAC CAATGCTACA GTTAAATTCA TCCTCACCGA CTCTTTTCCT TCTAACGTTG   2160
GTAATAGGTG GTTACGATGT CAATTTAAGT AGGAGTGGCT GAGAAAAGGA AGATTGCAAC
```

Fig. 6. (Continuation page 2, SEQ ID NO:6).

```
AAGTCGTTAT TTTGCAACCA TCTGCAGAAG AGCAGCACCA GATCGATATG TCCATTATAT    2220
TTCAGCAATA AAACGTTGGT AGACGTCTTC TCGTCGTGGT CTAGCTATAC AGGTAATATA

TCATTGCAGT GCTGGCTGGT GGTTGTGCTT TGCTACTTTT GGCCATCTTT TTTGTGGCCT    2280
AGTAACGTCA CGACCGACCA CCAACACGAA ACGATGAAAA CCGGTAGAAA AAACACCGGA

GTACTTGTAA AAAGAAAGCT GGTGAATTTA AGCAGGTACC TGAACAACAC GGAACATGCA    2340
CATGAACATT TTTCTTTCGA CCACTTAAAT TCGTCCATGG ACTTGTTGTG CCTTGTACGT

ATGAAGAACG CCTGTTAAGC ACCCCATCTC CCCAGTCGGT CTCTTCTTCT TTGTCTCAGT    2400
TACTTCTTGC GGACAATTCG TGGGGTAGAG GGGTCAGCCA GAGAAGAAGA AACAGAGTCA

CTGAGTCATG CCAACTCTCC ATCAATACTG AATCTGAGAA TTGCAGCGTG TCCTCTAACC    2460
GACTCAGTAC GGTTGAGAGG TAGTTATGAC TTAGACTCTT AACGTCGCAC AGGAGATTGG

AAGAGCAGCA TCAGCAAACA GGCATAAAGC ACTCCATCTC TGTACCATCT TATCACACAT    2520
TTCTCGTCGT AGTCGTTTGT CCGTATTTCG TGAGGTAGAG ACATGGTAGA ATAGTGTGTA

CTGGTTGGCA CCTGGACAAT TGTGCAATGA GCATAAGTGG ACATTCTCAC ATGGGGCACA    2580
GACCAACCGT GGACCTGTTA ACACGTTACT CGTATTCACC TGTAAGAGTG TACCCCGTGT

TTAGTACAAA GGTACAGTGG GCAAAGGAGA TAGTGACTTC AATGACAGTG ACTCTGATAC    2640
AATCATGTTT CCATGTCACC CGTTTCCTCT ATCACTGAAG TTACTGTCAC TGAGACTATG

TAGTGGAGAA TCAGAAAAGA AGAGCATTGA GCAGCCAATG CAGGCACAAG CCAGTGCTCA    2700
ATCACCTCTT AGTCTTTTCT TCTCGTAACT CGTCGGTTAC GTCCGTGTTC GGTCACGAGT

ATACACAGAT GAATCAGCAG GGTTCCGACA TGCCGATAAC TATTTCAGCC ACCGAATCAA    2760
TATGTGTCTA CTTAGTCGTC CCAAGGCTGT ACGGCTATTG ATAAAGTCGG TGGCTTAGTT

CAAGGGTCCA GAAAATGGGA ACTGCACATT GCAATATGAA AAGGGCTATA GACTGTCTTA    2820
GTTCCCAGGT CTTTTACCCT TGACGTGTAA CGTTATACTT TTCCCGATAT CTGACAGAAT

CTCTGTAGCT CCTGTATATT ACAATACCTA CCATGCAAGA ATGCCTAACC TGCACATACC    2880
GAGACATCGA GGACATATAA TGTTATGGAT GGTACGTTCT TACGGATTGG ACGTGTATGG

GAACCATACC CTTAGAGACC CTTATTACCA TATCAATAAT CCTGTTGCTA ATCGGATGCA    2940
CTTGGTATGG GAATCTCTGG GAATAATGGT ATAGTTATTA GGACAACGAT TAGCCTACGT

GGCGGAATAT GAAAGAGATT TAGTCAACAG AAGTGCAACG TTATCTCCGC AGAGATCGTC    3000
CCGCCTTATA CTTTCTCTAA ATCAGTTGTC TTCACGTTGC AATAGAGGCG TCTCTAGCAG

TAGCAGATAC CAAGAATTCA ATTACAGTCC GCAGATATCA AGACAGCTTC ATCCTTCAGA    3060
ATCGTCTATG GTTCTTAAGT TAATGTCAGG CGTCTATAGT TCTGTCGAAG TAGGAAGTCT

AATTGCTACA ACCTTTTAAT CATTAGGCAT GCAAGTGAGA ATGCACAAAG GCAAGTGCTT    3120
TTAACGATGT TGGAAAATTA GTAATCCGTA CGTTCACTCT TACGTGTTTC CGTTCACGAA

TAGCATGAAA GCTAAATATA TGGAGTCTCC CCTTTCCCTC TGATGGATGG GGGGAGACAC    3180
ATCGTACTTT CGATTTATAT ACCTCAGAGG GGAAAGGGAG ACTACCTACC CCCCTCTGTG

AGGACAGTGC ATAAATATAC AGCTGCTTTC TATTTGCATT TCACTTGGGA ATTTTTTGTT    3240
TCCTGTCACG TATTTATATG TCGACGAAAG ATAAACGTAA AGTGAACCCT TAAAAAACAA

TTTTTTACAT ATTTATTTTT CCTGAATTGA ATGTGACATT GTCCTGTCAC CTAACTAGCA    3300
AAAAAATGTA TAAATAAAAA GGACTTAACT TACACTGTAA CAGGACAGTG GATTGATCGT
```

Fig. 6. (Continuation page 3, SEQ ID NO:6).

ATTAAATCCA CAGACCTACA GTCAAATATT TGAGGGCCCC TGAAACAGCA CATCAGTCAG 3360
TAATTTAGGT GTCTGGATGT CAGTTTATAA ACTCCCGGGG ACTTTGTCGT GTAGTCAGTC

GACCTAAAGT GGCCTTTTTA CTTTTAGCAG CTCCTGGGTC TGCCCTCTGT GTTAATCAGC 3420
CTGGATTTCA CCGGAAAAAT GAAAATCGTC GAGGACCCAG ACGGGAGACA CAATTAGTCG

CCCTGGTCAA GTCCTGAGTA GGATCATGGC GTTTTTATAT GCATCTCACC TACTTTGGAC 3480
GGGACCAGTT CAGGACTCAT CCTAGTACCG CAAAAATATA CGTAGAGTGG ATGAAACCTG

GTGATTTACA CATAATAGGA AACGCTTGGT TTCAGTGAAG TCTGTGTTGT ATATATTCTG 3540
CACTAAATGT GTATTATCCT TTGCGAACCA AAGTCACTTC AGACACAACA TATATAAGAC

TTATATACAC GCATTTTGTG TTTGTGTATA TATTTCAAGT CCATTCAGAT ATGTGTATAT 3600
AATATATGTG CGTAAAACAC AAACACATAT ATAAAGTTCA GGTAAGTCTA TACACATATA

AGTGCAGACC TTGTAAATTA AATATTCTGA TACTTTTTCC TCAATAAATA TTTAAAT
TCACGTCTGG AACATTTAAT TTATAAGACT ATGAAAAAGG AGTTATTTAT AAATTTA

Fig. 6. (Continuation page 4, SEQ ID NO:6).

MVCCGPGRML LGWAGLLVLA ALCLLQVPGA QAAACEPVRI PLCKSLPWNM TKMPNHLHHS 60

TQANAILAME QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPCKS VCERARQGCE 120

PILIKYRHSW PESLACDELP VYDRGVCISP EAIVTADGAD FPMDSSTGHC RGASSERCKC 180

KPVRATQKTY FRNNYNYVIR AKVKEVKMKC HDVTAVVEVK EILKASLVNI PRDTVNLYTT 240

SGCLCPPLTV NEEYVIMGYE DEERSRLLLV EGSIAEKWKD RLGKKVKRWD MKLRHLGLGK 300

TDASDSTQNQ KSGRNSNPRP ARS.

Figure 7. Deduced amino acid sequence of mouse FRZB-1 protein. SEQ ID NO:7.

Figure 8. Nucleotide sequence of the full-length mouse FRZB-1 cDNA. SEQ ID NO:8.

```
AAGCCTGGGA CCATGGTCTG CTGCGGCCCG GGACGGATGC TGCTAGGATG GGCCGGGTTG    60
TTCGGACCCT GGTACCAGAC GACGCCGGGC CCTGCCTACG ACGATCCTAC CCGGCCCAAC

CTAGTCCTGG CTGCTCTCTG CCTGCTCCAG GTGCCCGGAG CTCAGGCTGC AGCCTGTGAG   120
GATCAGGACC GACGAGAGAC GGACGAGGTC CACGGGCCTC GAGTCCGACG TCGGACACTC

CCTGTCCGCA TCCCGCTGTG CAAGTCCCTT CCCTGGAACA TGACCAAGAT GCCCAACCAC   180
GGACAGGCGT AGGGCGACAC GTTCAGGGAA GGGACCTTGT ACTGGTTCTA CGGGTTGGTG

CTGCACCACA GCACCCAGGC TAACGCCATC CTGGCCATGG AACAGTTCGA AGGGCTGCTG   240
GACGTGGTGT CGTGGGTCCG ATTGCGGTAG GACCGGTACC TTGTCAAGCT TCCCGACGAC

GGCACCCACT GCAGCCCGGA TCTTCTCTTC TTCCTCTGTG CAATGTACGC ACCCATTTGC   300
CCGTGGGTGA CGTCGGGCCT AGAAGAGAAG AAGGAGACAC GTTACATGCG TGGGTAAACG

ACCATCGACT TCCAGCACGA GCCCATCAAG CCCTGCAAGT CTGTGTGTGA GCGCGCCCGA   360
TGGTAGCTGA AGGTCGTGCT CGGGTAGTTC GGGACGTTCA GACACACACT CGCGCGGGCT

CAGGGCTGCG AGCCCATTCT CATCAAGTAC CGCCACTCGT GGCCGGAAAG CTTGGCCTGC   420
GTCCCGACGC TCGGGTAAGA GTAGTTCATG GCGGTGAGCA CCGGCCTTTC GAACCGGACG

GACGAGCTGC CGGTGTACGA CCGCGGCGTG TGCATCTCTC CTGAGGCCAT CGTCACCGCG   480
CTGCTCGACG GCCACATGCT GGCGCCGCAC ACGTAGAGAG GACTCCGGTA GCAGTGGCGC

GACGGAGCGG ATTTTCCTAT GGATTCAAGT ACTGGACACT GCAGAGGGGC AAGCAGCGAA   540
CTGCCTCGCC TAAAAGGATA CCTAAGTTCA TGACCTGTGA CGTCTCCCCG TTCGTCGCTT

CGTTGCAAAT GTAAGCCTGT CAGAGCTACA CAGAAGACCT ATTTCCGGAA CAATTACAAC   600
GCAACGTTTA CATTCGGACA GTCTCGATGT GTCTTCTGGA TAAAGGCCTT GTTAATGTTG

TATGTCATCC GGGCTAAAGT TAAAGAGGTA AAGATGAAAT GTCATGATGT GACCGCCGTT   660
ATACAGTAGG CCCGATTTCA ATTTCTCCAT TTCTACTTTA CAGTACTACA CTGGCGGCAA

GTGGAAGTGA AGGAAATTCT AAAGGCATCA CTGGTAAACA TTCCAAGGGA CACCGTCAAT   720
CACCTTCACT TCCTTTAAGA TTTCCGTAGT GACCATTTGT AAGGTTCCCT GTGGCAGTTA

CTTTATACCA CCTCTGGCTG CCTCTGTCCT CCACTTACTG TCAATGAGGA ATATGTCATC   780
GAAATATGGT GGAGACCGAC GGAGACAGGA GGTGAATGAC AGTTACTCCT TATACAGTAG

ATGGGCTATG AAGACGAGGA ACGTTCCAGG TTACTCTTGG TAGAAGGCTC TATAGCTGAG   840
TACCCGATAC TTCTGCTCCT TGCAAGGTCC AATGAGAACC ATCTTCCGAG ATATCGACTC

AAGTGGAAGG ATCGGCTTGG TAAGAAAGTC AAGCGCTGGG ATATGAAACT CCGACACCTT   900
TTCACCTTCC TAGCCGAACC ATTCTTTCAG TTCGCGACCC TATACTTTGA GGCTGTGGAA

GGACTGGGTA AAACTGATGC TAGCGATTCC ACTCAGAATC AGAAGTCTGG CAGGAACTCT   960
CCTGACCCAT TTTGACTACG ATCGCTAAGG TGAGTCTTAG TCTTCAGACC GTCCTTGAGA
```

AATCCCCGGC CAGCACGCAG CTAAATCCTG AAATGTAAAA GGCCACACCC ACGGACTCCC 1020
TTAGGGGCCG GTCGTGCGTC GATTTAGGAC TTTACATTTT CCGGTGTGGG TGCCTGAGGG

TTCTAAGACT GGCGCTGGTG GACTAACAAA GGAAAACCGC ACAGTTGTGC TCGTGACCGA 1080
AAGATTCTGA CCGCGACCAC CTGATTGTTT CCTTTTGGCG TGTCAACACG AGCACTGGCT

TTGTTTACCG CAGACACCGC GTGGCTACCG AAGTTACTTC CGGTCCCCTT TCTCCTGCTT 1140
AACAAATGGC GTCTGTGGCG CACCGATGGC TTCAATGAAG GCCAGGGGAA AGAGGACGAA

CTTAATGGCG TGGGGTTAGA TCCTTTAATA TGTTATATAT TCTGTTTCAT CAATCACGTG 1200
GAATTACCGC ACCCCAATCT AGGAAATTAT ACAATATATA AGACAAAGTA GTTAGTGCAC

GGGACTGTTC TTTTGCAACC AGAATAGTAA ATTAAATATG TTGATGCTAA GGTTTCTGTA 1260
CCCTGACAAG AAAACGTTGG TCTTATCATT TAATTTATAC AACTACGATT CCAAAGACAT

CTGGACTCCC TGGGTTTAAT TTGGTGTTCT GTACCCTGAT TGAGAATGCA ATGTTTCATG 1320
GACCTGAGGG ACCCAAATTA AACCACAAGA CATGGGACTA ACTCTTACGT TACAAAGTAC

TAAAGAGAGA ATCCTGGTCA TATCTCAAGA ACTAGATATT GCTGTAAGAC AGCCTCTGCT 1380
ATTTCTCTCT TAGGACCAGT ATAGAGTTCT TGATCTATAA CGACATTCTG TCGGAGACGA

GCTGCGCTTA TAGTCTTGTG TTTGTATGCC TTTGTCCATT TCCCTCATGC TGTGAAAGTT 1440
CGACGCGAAT ATCAGAACAC AAACATACGG AAACAGGTAA AGGGAGTACG ACACTTTCAA

ATACATGTTT ATAAAGGTAG AACGGCATTT TGAAATCAGA CACTGCACAA GCAGAGTAGC 1500
TATGTACAAA TATTTCCATC TTGCCGTAAA ACTTTAGTCT GTGACGTGTT CGTCTCATCG

CCAACACCAG GAAGCATTTA TGAGGAAACG CCACACAGCA TGACTTATTT TCAAGATTGG 1560
GGTTGTGGTC CTTCGTAAAT ACTCCTTTGC GGTGTGTCGT ACTGAATAAA AGTTCTAACC

CAGGCAGCAA AATAAATAGT GTTGGGAGCC AAGAAAAGAA TATTTTGCCT GGTTAAGGGG 1620
GTCCGTCGTT TTATTTATCA CAACCCTCGG TTCTTTTCTT ATAAAACGGA CCAATTCCCC

CACACTGGAA TCAGTAGCCC TTGAGCCATT AACAGCAGTG TTCTTCTGGC AAGTTTTTGA 1680
GTGTGACCTT AGTCATCGGG AACTCGGTAA TTGTCGTCAC AAGAAGACCG TTCAAAAACT

TTTGTTCATA AATGTATTCA CGAGCATTAG AGATGAACTT ATAACTAGAC ATCTGTTGTT 1740
AAACAAGTAT TTACATAAGT GCTCGTAATC TCTACTTGAA TATTGATCTG TAGACAACAA

ATCTCTATAG CTCTGCTTCC TTCTAAATCA AACCCATTGT TGGATGCTCC CTCTCCATTC 1800
TAGAGATATC GAGACGAAGG AAGATTTAGT TTGGGTAACA ACCTACGAGG GAGAGGTAAG

ATAAATAAAT TTGGCTTGCT GTATTGGCCA GGAAAAGAAA GTATTAAAGT ATGCATGCAT 1860
TATTTATTTA AACCGAACGA CATAACCGGT CCTTTTCTTT CATAATTTCA TACGTACGTA

GTGCACCAGG GTGTTATTTA ACAGAGGTAT GTAACTCTAT AAAAGACTAT AATTTACAGG 1920
CACGTGGTCC CACAATAAAT TGTCTCCATA CATTGAGATA TTTTCTGATA TTAAATGTCC

ACACGGAAAT GTGCACATTT GTTTACTTTT TTTCTTCCTT TTGCTTTGGG CTTGTGATTT 1980
TGTGCCTTTA CACGTGTAAA CAAATGAAAA AAAGAAGGAA AACGAAACCC GAACACTAAA

TGGTTTTTGG TGTGTTTATG TCTGTATTTT GGGGGGTGGG TAGGTTTAAG CCATTGCACA 2040
ACCAAAAACC ACACAAATAC AGACATAAAA CCCCCCACCC ATCCAAATTC GGTAACGTGT

TTCAAGTTGA ACTAGATTAG AGTAGACTAG GCTCATTGGC CTAGACATTA TGATTTGAAT 2100
AAGTTCAACT TGATCTAATC TCATCTGATC CGAGTAACCG GATCTGTAAT ACTAAACTTA

TTGTGTTGTT TAATGCTCCA TCAAGATGTC TAATAAAAGG AATATGGTTG TCAACAGAGA 2160
AACACAACAA ATTACGAGGT AGTTCTACAG ATTATTTTCC TTATACCAAC AGTTGTCTCT

CGACAACAAC AACAAA
GCTGTTGTTG TTGTTT

```
MVCGSPGGML LLRAGLLALA ALCLLRVPGA RAAACEPVRI PLCKSLPWNM TKMPNHLHHS    60
TQANAILAIE QFEGLLGTHC SPDLLFFLCA MYAPICTIDF QHEPIKPCKS VCERARQGCE   120
PILIKYRHSW PENLACEELP VYDRGVCISP EAIVTADGAD FPMDSSNGNC RGASSERCKC   180
KPIRATQKTY FRNNYNYVIR AKVKEIKTKC HDVTAVVEVK EILKSSLVNI PRDTVNLYTS   240
SGCLCPPLNV NEEYIIMGYE DEERSRLLLV EGSIAEKWKD RLGKKVKRWD MKLRHLGLSK   300
SDSSNSDSTQ SQKSGRNSNP RQARN.
```

Figure 9. Deduced amino acid sequence of human FRZB-1 protein. SEQ ID NO:9.

Figure 10. Nucleotide sequence of the full-length human FRZB-1 cDNA. SEQ ID NO:10. This sequence was assembled from public ESTs from the Genbank database (accession numbers: H18848, R63748, W38677, W44760, H38379 and N71244).

```
GGCGGAGCGG GCCTTTTGGC GTCCACTGCG CGGCTGCACC CTGCCCCATC TGCCGGGATC    60
CCGCCTCGCC CGGAAAACCG CAGGTGACGC GCCGACGTGG GACGGGGTAG ACGGCCCTAG

ATGGTCTGCG GCAGCCCGGG AGGGATGCTG CTGCTGCGGG CCGGGCTGCT TGCCCTGGCT   120
TACCAGACGC CGTCGGGCCC TCCCTACGAC GACGACGCCC GGCCCGACGA ACGGGACCGA

GCTCTCTGCC TGCTCCGGGT GCCCGGGGCT CGGGCTGCAG CCTGTGAGCC CGTCCGCATC   180
CGAGAGACGG ACGAGGCCCA CGGGCCCCGA GCCCGACGTC GGACACTCGG GCAGGCGTAG

CCCCTGTGCA AGTCCCTGCC CTGGAACATG ACTAAGATGC CCAACCACCT GCACCACAGC   240
GGGGACACGT TCAGGGACGG GACCTTGTAC TGATTCTACG GGTTGGTGGA CGTGGTGTCG

ACTCAGGCCA ACGCCATCCT GGCCATCGAG CAGTTCGAAG GTCTGCTGGG CACCCACTGC   300
TGAGTCCGGT TGCGGTAGGA CCGGTAGCTC GTCAAGCTTC CAGACGACCC GTGGGTGACG

AGCCCCGATC TGCTCTTCTT CCTCTGTGCC ATGTACGCGC CCATCTGCAC CATTGACTTC   360
TCGGGGCTAG ACGAGAAGAA GGAGACACGG TACATGCGCG GGTAGACGTG GTAACTGAAG

CAGCACGAGC CCATCAAGCC CTGTAAGTCT GTGTGCGAGC GGGCCCGGCA GGGCTGTGAG   420
GTCGTGCTCG GGTAGTTCGG GACATTCAGA CACACGCTCG CCCGGGCCGT CCCGACACTC

CCCATACTCA TCAAGTACCG CCACTCGTGG CCGGAGAACC TGGCCTGCGA GGAGCTGCCA   480
GGGTATGAGT AGTTCATGGC GGTGAGCACC GGCCTCTTGG ACCGGACGCT CCTCGACGGT

GTGTACGACA GGGGCGTGTG CATCTCTCCC GAGGCCATCG TTACTGCGGA CGGAGCTGAT   540
CACATGCTGT CCCCGCACAC GTAGAGAGGG CTCCGGTAGC AATGACGCCT GCCTCGACTA

TTTCCTATGG ATTCTAGTAA CGGAAACTGT AGAGGGGCAA GCAGTGAACG CTGTAAATGT   600
AAAGGATACC TAAGATCATT GCCTTTGACA TCTCCCCGTT CGTCACTTGC GACATTTACA

AAGCCTATTA GAGCTACACA GAAGACCTAT TTCCGGAACA ATTACAACTA TGTCATTCGG   660
TTCGGATAAT CTCGATGTGT CTTCTGGATA AAGGCCTTGT TAATGTTGAT ACAGTAAGCC

GCTAAAGTTA AAGAGATAAA GACTAAGTGC CATGATGTGA CTGCAGTAGT GGAGGTGAAG   720
CGATTTCAAT TTCTCTATTT CTGATTCACG GTACTACACT GACGTCATCA CCTCCACTTC

GAGATTCTAA AGTCCTCTCT GGTAAACATT CCACGGGACA CTGTCAACCT CTATACCAGC   780
CTCTAAGATT TCAGGAGAGA CCATTTGTAA GGTGCCCTGT GACAGTTGGA GATATGGTCG

TCTGGCTGCC TCTGCCCTCC ACTTAATGTT AATGAGGAAT ATATCATCAT GGGCTATGAA   840
AGACCGACGG AGACGGGAGG TGAATTACAA TTACTCCTTA TATAGTAGTA CCCGATACTT
```

GATGAGGAAC GTTCCAGATT ACTCTTGGTG GAAGGCTCTA TAGCTGAGAA GTGGAAGGAT 900
CTACTCCTTG CAAGGTCTAA TGAGAACCAC CTTCCGAGAT ATCGACTCTT CACCTTCCTA

CGACTCGGTA AAAAAGTTAA GCGCTGGGAT ATGAAGCTTC GTCATCTTGG ACTCAGTAAA 960
GCTGAGCCAT TTTTTCAATT CGCGACCCTA TACTTCGAAG CAGTAGAACC TGAGTCATTT

AGTGATTCTA GCAATAGTGA TTCCACTCAG AGTCAGAAGT CTGGCAGGAA CTCGAACCCC 1020
TCACTAAGAT CGTTATCACT AAGGTGAGTC TCAGTCTTCA GACCGTCCTT GAGCTTGGGG

CGGCAAGCAC GCAACTAAAT CCCGAAATAC AAAAAGTAAC ACAGTGGACT TCCTATTAAG 1080
GCCGTTCGTG CGTTGATTTA GGGCTTTATG TTTTTCATTG TGTCACCTGA AGGATAATTC

ACTTACTTGC ATTGCTGGAC TAGCAAAGGA AAATTGCACT ATTGCACATC ATATTCTATT 1140
TGAATGAACG TAACGACCTG ATCGTTTCCT TTTAACGTGA TAACGTGTAG TATAAGATAA

GTTTACTATA AAAATCATGT GATAACTGAT TATTACTTCT GTTTCTCTTT TGGTTTCTGC 1200
CAAATGATAT TTTTAGTACA CTATTGACTA ATAATGAAGA CAAAGAGAAA ACCAAAGACG

TTCTCTCTTC TCTCAACCCC TTTGTAATGG TTTGGGGGCA GACTCTTAAG TATATTGTGA 1260
AAGAGAGAAG AGAGTTGGGG AAACATTACC AAACCCCCGT CTGAGAATTC ATATAACACT

GTTTTCTATT TCACTAATCA TGAGAAAAAC TGTTCTTTTG CAATAATAAT AAATTAAACA 1320
CAAAAGATAA AGTGATTAGT ACTCTTTTTG ACAAGAAAAC GTTATTATTA TTTAATTTGT

TGCTGTTACC AGAGCCTCTT TGCTGAGTCT CCAGATGTTA ATTTACTTTC TGCACCCCAA 1380
ACGACAATGG TCTCGGAGAA ACGACTCAGA GGTCTACAAT TAAATGAAAG ACGTGGGGTT

TTGGGAATGC AATATTGGAT GAAAAGAGAG GTTTCTGGTA TTCACAGAAA GCTAGATATG 1440
AACCCTTACG TTATAACCTA CTTTTCTCTC CAAAGACCAT AAGTGTCTTT CGATCTATAC

CCTTAAAACA TACTCTGCCG ATCTAATTAC AGCCTTATTT TTGTATGCCT TTTGGGCATT 1500
GGAATTTTGT ATGAGACGGC TAGATTAATG TCGGAATAAA AACATACGGA AAACCCGTAA

CTCCTCATGC TTAGAAAGTT CCAAATGTTT ATAAAGGTAA AATGGCAGTT TGAAGTCAAA 1560
GAGGAGTACG AATCTTTCAA GGTTTACAAA TATTTCCATT TTACCGTCAA ACTTCAGTTT

TGTCACATAG GCAAAGCAAT CAAGCACCAG GAAGTGTTTA TGAGGAAACA ACACCCAAGA 1620
ACAGTGTATC CGTTTCGTTA GTTCGTGGTC CTTCACAAAT ACTCCTTTGT TGTGGGTTCT

TGAATTATTT TTGAGACTGT CAGGAAGTAA AATAAATAGG AGCTTAAGAA AGAACATTTT 1680
ACTTAATAAA AACTCTGACA GTCCTTCATT TTATTTATCC TCGAATTCTT TCTTGTAAAA

GCCTGATTGA GAAGCACAAC TGAAACCAGT AGCCGCTGGG GTGTTAATGG TAGCATTCTT 1740
CGGACTAACT CTTCGTGTTG ACTTTGGTCA TCGGCGACCC CACAATTACC ATCGTAAGAA

CTTTTGGCAA TACATTTGAT TTGTTCATGA ATATATTAAT CAGCATTAGA GAAATGAATT 1800
GAAAACCGTT ATGTAAACTA AACAAGTACT TATATAATTA GTCGTAATCT CTTTACTTAA

ATAACTAGAC ATCTGCTGTT ATCACCATAG TTTTGTTTAA TTTGCTTCCT TTTAAATAAA 1860
TATTGATCTG TAGACGACAA TAGTGGTATC AAAACAAATT AAACGAAGGA AAATTTATTT

CCCATTGGTG AAAGTCAAAA AAAAAAAAAA AAA
GGGTAACCAC TTTCAGTTTT TTTTTTTTTT TTT